

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
17 juillet 2003 (17.07.2003)

PCT

(10) Numéro de publication internationale
WO 2003/057648 A3

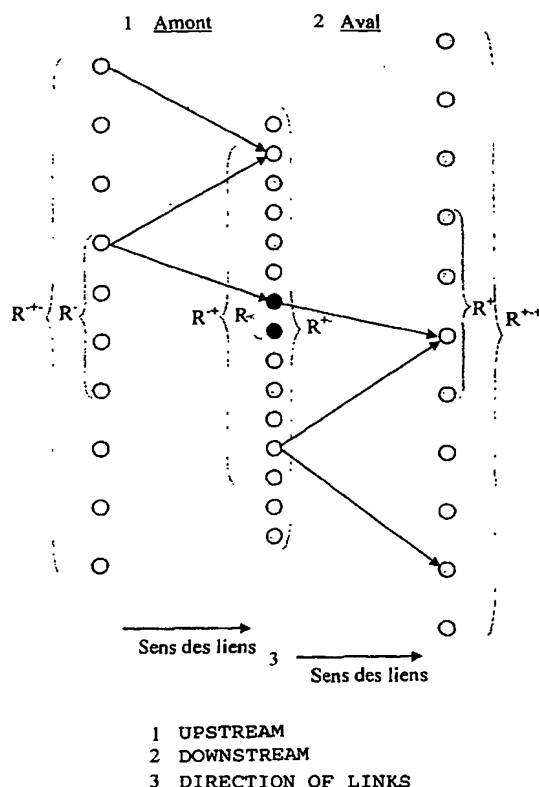
(51) Classification internationale des brevets⁷ : G06F 17/30
(21) Numéro de la demande internationale :
PCT/FR2003/000089
(22) Date de dépôt international :
13 janvier 2003 (13.01.2003)
(25) Langue de dépôt : français
(26) Langue de publication : français
(30) Données relatives à la priorité :
02/00341 11 janvier 2002 (11.01.2002) FR
02/05751 7 mai 2002 (07.05.2002) FR

(71) Déposant et
(72) Inventeur : MAIM, Enrico [FR/FR]; 17, rue Biscornet,
F-75012 Paris (FR).
(74) Mandataires : MARTIN, Jean-Jacques etc.; Cabinet
Régimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17
(FR).
(81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,
DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG,
SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC,
VN, YU, ZA, ZM, ZW.

[Suite sur la page suivante]

(54) Title: METHODS AND SYSTEMS FOR SEARCHING AND ASSOCIATING INFORMATION RESOURCES SUCH AS
WEB PAGES

(54) Titre : PROCÉDES ET SYSTEMES DE RECHERCHE ET D'ASSOCIATION DE RESSOURCES D'INFORMATION
TELLES QUE DES PAGES WEB



(57) Abstract: Disclosed is a method for detecting additional relevant resources in relation to a predefined initial number of resources, consisting of the following steps: a) a set of referencing resources consisting of all resources being linked to at least one of the initial resources is detected; b) a set of potential resources consisting of all resources referenced by the referencing resources is established; c) for each potential resource, a potential-resource relevance score is calculated between said potential resource and all initial resources on the basis of existing links in the referencing resources to the potential resource and the initial resources, and also on the basis of the relevance scores of referencing resources assigned to each referencing resource; d) for each referencing resource, a referencing-resource relevance score is recalculated on the basis of links occurring in said referencing resource to the potential resources and also on the basis of the potential-resource relevance score assigned to the potential resources in step c); e) step c), and if necessary several times step d) followed by step c), is repeated if necessary; f) said additional relevant resources are identified as the potential resources having the highest potential-resource relevance scores.

(57) Abrégé: Un procédé pour déterminer des ressources additionnelles pertinentes par rapport à un ensemble donné de ressources de départ comprend les étapes suivantes : a) identifier un ensemble de ressources citantes constituées par toutes les ressources ayant un lien vers au moins l'une des ressources de départ, b) former un ensemble de ressources candidates constitué par l'ensemble des ressources citées par les ressources citantes, c) pour chaque ressource candidate, calculer un score de pertinence de ressource candidate entre ladite ressource candidate et l'ensemble de ressources de départ

[Suite sur la page suivante]

WO 2003/057648 A3



(84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(88) Date de publication du rapport de recherche internationale:

17 juin 2004

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

Publiée :

— avec rapport de recherche internationale

sur la base de l'existence de liens situés dans les ressources citantes et dirigés vers la ressource candidate et vers les ressources de départ, et sur la base également de scores de pertinence de ressources citantes affectés à chacune des ressources citantes, d) pour chaque ressource citante, recalculer un score de pertinence de ressource citante sur la base de l'existence, dans la ressource citante en question, de liens vers les ressources candidates et sur la base également des scores de pertinence de ressource candidate attribués aux ressources candidates à l'étape c), e) répéter le cas échéant l'étape c) et le cas échéant une ou plusieurs fois l'étape d) puis l'étape c), f) déterminer lesdites ressources additionnelles pertinentes comme étant les ressources candidates qui présentent les meilleurs scores de pertinence de ressource candidate.

INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 03/00089

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DEAN J ET AL: "Finding related pages in the World Wide Web" COMPUTER NETWORKS, ELSEVIER SCIENCE PUBLISHERS B.V., AMSTERDAM, NL, vol. 31, no. 11-16, 17 May 1999 (1999-05-17), pages 1467-1479, XP004304567 ISSN: 1389-1286 page 1467, right-hand column, line 11 - page 1468, right-hand column, line 4 page 1469, left-hand column, paragraph 2 - page 1471, right-hand column, paragraph 2.3 page 1472, left-hand column, line 41 - last line page 1478, right-hand column, line 10 - line 13</p> <p style="text-align: center;">----- -/--</p>	1-4

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

11 August 2003

Date of mailing of the international search report

12 11 2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Fournier, C

INTERNATIONAL SEARCH REPORT

International Application No
PCT/TR 03/00089

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>PIROLI P ET AL: "SILK FROM A SOW'S EAR: EXTRACTING USABLE STRUCTURES FROM THE WEB" COMMON GROUND. CHI '96 CONFERENCE PROCEEDINGS. CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS. VANCOUVER, APRIL 13 - 18, 1996, CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, NEW YORK, ACM, US, 13 April 1996 (1996-04-13), pages 118-125, XP000657810 ISBN: 0-201-94687-4 page 118, left-hand column, last paragraph - right-hand column, line 13</p> <p>-----</p>	1,4
A	<p>BICHTLER J ET AL: "THE COMBINED USE OF BIBLIOGRAPHIC COUPLING AND COCITATION FOR DOCUMENT RETRIEVAL" JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, AMERICAN SOCIETY FOR INFORMATION. WASHINGTON, US, vol. 31, no. 4, 1 July 1980 (1980-07-01), pages 278-282, XP002043307 ISSN: 0002-8231 abstract page 278, left-hand column, line 13 - right-hand column, line 1 page 278, right-hand column, last paragraph - page 279, left-hand column, line 4</p> <p>-----</p>	1,4

The International Searching Authority has determined that this international application contains more than one invention or group of inventions, namely:

1. Claims 1-4

Method for determining additional relevant resources relative to a given starting set of resources.

2. Claims 5-13

Methods and system for finding fresh information resources on the basis of existing information resources.

3. Claims 14-24

Method for managing information resources in a computer system using a graphic interface.

4. Claims 25-27

Method for identifying a part of a resource that is capable of constituting a significant search engine query.

5. Claims 28-36

Method for managing information resources in a computer system by resource association and suggestion.

6. Claim 37

Method for identifying recent information resources that are accessible via links and have a content similar to that of a given resource.

7. Claims 38-49

Method for enabling a user to access relevant information entities from a starting information entity.

8. Claims 50-52

Methods for determining the relevance scores assigned to text units in a text document.

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No
PCT/FR 03/00089

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 G06F17/30

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 7 G06F

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, INSPEC, WPI Data, PAJ

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X .	<p>DEAN J ET AL: "Finding related pages in the World Wide Web"</p> <p>COMPUTER NETWORKS, ELSEVIER SCIENCE PUBLISHERS B.V., AMSTERDAM, NL, vol. 31, no. 11-16, 17 mai 1999 (1999-05-17), pages 1467-1479, XP004304567</p> <p>ISSN: 1389-1286</p> <p>page 1467, colonne de droite, ligne 11 -</p> <p>page 1468, colonne de droite, ligne 4</p> <p>page 1469, colonne de gauche, alinéa 2 -</p> <p>page 1471, colonne de droite, alinéa 2.3</p> <p>page 1472, colonne de gauche, ligne 41 - dernière ligne</p> <p>page 1478, colonne de droite, ligne 10 - ligne 13</p> <p>-----</p> <p>-/--</p>	1-4



Voir la suite du cadre C pour la fin de la liste des documents



Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

- "T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
- "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
- "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
- "&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

11 août 2003

Date d'expédition du présent rapport de recherche internationale

12. 11. 2003

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Fournier, C.

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No
PCT/FR 03/00089

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	<p>PIROLI P ET AL: "SILK FROM A SOW'S EAR: EXTRACTING USABLE STRUCTURES FROM THE WEB" COMMON GROUND. CHI '96 CONFERENCE PROCEEDINGS. CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS. VANCOUVER, APRIL 13 - 18, 1996, CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, NEW YORK, ACM, US, 13 avril 1996 (1996-04-13), pages 118-125, XP000657810 ISBN: 0-201-94687-4 page 118, colonne de gauche, dernier alinéa - colonne de droite, ligne 13</p>	1,4
A	<p>BICHTLER J ET AL: "THE COMBINED USE OF BIBLIOGRAPHIC COUPLING AND COCITATION FOR DOCUMENT RETRIEVAL" JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE, AMERICAN SOCIETY FOR INFORMATION. WASHINGTON, US, vol. 31, no. 4, 1 juillet 1980 (1980-07-01), pages 278-282, XP002043307 ISSN: 0002-8231 abrégé page 278, colonne de gauche, ligne 13 - colonne de droite, ligne 1 page 278, colonne de droite, dernier alinéa - page 279, colonne de gauche, ligne 4</p>	1,4

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR 03/00089

Cadre I Observations - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 1 de la première feuille)

Conformément à l'article 17.2)a), certaines revendications n'ont pas fait l'objet d'une recherche pour les motifs suivants:

1. ☐ Les revendications n^{os} se rapportent à un objet à l'égard duquel l'administration n'est pas tenue de procéder à la recherche, à savoir:
2. ☐ Les revendications n^{os} se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectuée, en particulier:
3. ☐ Les revendications n^{os} sont des revendications dépendantes et ne sont pas rédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

Cadre II Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 2 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☐ Comme toutes les taxes additionnelles ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.
2. ☐ Comme toutes les recherches portant sur les revendications qui s'y prêtaient ont pu être effectuées sans effort particulier justifiant une taxe additionnelle, l'administration n'a sollicité le paiement d'aucune taxe de cette nature.
3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n^{os}
4. ☒ Aucune taxe additionnelle demandée n'a été payée dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n^{os}
1-4

Remarque quant à la réserve

- ☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant.
- ☐ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: 1-4

Procédé pour déterminer des ressources additionnelles pertinentes par rapport à un ensemble donné de ressources de départ

2. revendications: 5-13

Procédés et système de recherche de nouvelles ressources d'information à partir de ressources d'informations existantes

3. revendications: 14-24

Procédé de gestion de ressources d'informations dans un système informatique à l'aide d'une interface graphique

4. revendications: 25-27

Procédé pour identifier une partie d'une ressource susceptible de constituer une requête significative pour un moteur de recherche

5. revendications: 28-36

Procédé de gestion de ressources d'informations dans un système informatique par association et suggestion de ressources

6. revendication: 37

Procédé pour identifier des ressources d'informations récentes accessibles par liens et ayant un contenu similaire à une ressource donnée

7. revendications: 38-49

Procédé pour permettre l'accès par un utilisateur à des entités d'informations pertinentes à partir d'une entité d'informations de départ

8. revendications: 50-52

SUITE DES RENSEIGNEMENTS INDICUES SUR PCT/ISA/ 210

Procédés pour déterminer des scores de pertinences attribués
à des unités de texte dans un document textuel
